Project Weekly report

Topic:- Modelling and Prediction of Athletic Readiness based on Training Load

Group Name: QuadML

Project Definition: 7

Group Member's names:-

- **Shashvi Shah, Riya Sudani, Dev Dave , Jeel Patel, Yashaswi Patel**

Feature Selection

- Our dataset includes athlete performance metrics, training loads, heart rate data, and other physical activity measures. During feature selection, the following approaches were applied specifically to your data:

  - Correlation Analysis:

    - For example, we examined the correlation between features like "Body Weight," "Peak Power," "Training Load Score," and target variables such as "Recovery Time" or "Performance Metrics."
    - Features with high multicollinearity (e.g., "HR.avg.." and "HR.max.." if strongly correlated) were considered for removal to avoid redundancy.
  - Handling Categorical Data:
    - "Session Name," "Type," and "Phase Name" were encoded using one-hot encoding to convert categorical labels into numerical format.
    - This allowed models to interpret categorical distinctions such as "Training" vs. "Game" sessions.
  - Scaling and Normalization:
    - Since your dataset includes a mix of large-scale metrics (e.g., "Total Distance" in meters) and smaller ones (e.g., "Jump Height"), normalization was necessary.
    - Min-Max scaling helped bring all features into a similar range, enhancing model performance.

- Next week
  - Start training initial models.